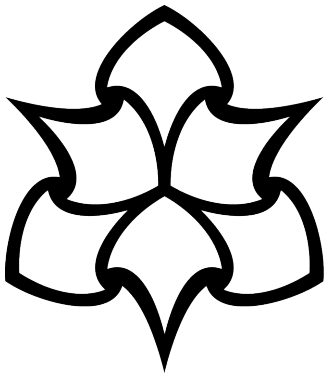

Why do some people search for their files much more than others? A preliminary study

Ofer Bergman
Department of Information Science, Bar-Ilan University, Ramat Gan, Israel
Tamar Israeli
Information Science, Bar-Ilan University, Ramat Gan, Israel, and
Yael Benn
Psychology, Manchester Metropolitan University, Manchester, UK

Received 12 August 2020
Revised 10 December 2020
4 February 2021
Accepted 21 March 2021

Abstract

Purpose – Previous research has repeatedly shown that people only search for files in a small minority of cases when they do not remember the file's location. The current study aimed to examine whether there is a group of hyper-searchers who search significantly more than others. Based on previous neurocognitive studies, this study aims to hypothesize that if such a group exists, they will have superior verbal memory and reduced visuospatial memory.

Design/methodology/approach – In total, 65 participants completed a questionnaire estimating their search percentages, as well as reporting demographic data. Verbal memory was measured using the Wechsler logical memory test, and visuospatial memory was assessed using an online card memory game.

Findings – Hyper-searchers were defined as participants with search percentage of over one standard deviation (SD) above the mean. The average search percentage of the seven participants who met this criterion was 51% (SD = 14%), over five times more than the other participants ($M = 10\%$, SD = 9%). Similar results were obtained by re-analyzing data from four previous papers ($N = 1{,}252$). The results further confirmed the hypothesis that hyper-searchers have significantly better verbal memory than other participants, possibly making searching easier and more successful for them. Lastly, the search percentage was positively predicted by verbal memory scores and negatively predicted by visuospatial memory scores. Explanations and future research are discussed.

Originality/value – This preliminary study is the first to introduce the concept of hyper-searchers, demonstrate its existence and study its causes.

Keywords Search percentage, File retrieval, Navigation, Verbal memory, Visuospatial memory

Paper type Research paper

Introduction

File retrieval is a basic activity performed by billions of computer users several times a day. The two main ways in which files can be retrieved are *navigation* and *search*. When navigating, users manually traverse their folder hierarchy until they reach the one in which the target item is stored. For search, users first generate a query specifying an attribute of the target item, and when the search engine returns a set of results, they select the relevant item from the list (Bergman *et al.*, 2008). Although search is more flexible than navigation, and despite advances in search engine technology, file retrieval studies consistently show navigation preference over search (Barreau and Nardi, 1995; Boardman and Sasse, 2004; Fitchett and Cockburn, 2015; Bergman *et al.*, 2014). Furthermore, participants search only as a last resort, in the minority of cases in which they forget the folder where they stored the target file (Bergman *et al.*, 2008). However, while this is true for the large majority of participants, we noticed that the search frequency is a skewed distribution with a long tail, such that a minority of people tend to perform many more searches compared to others. We term these people *hyper-searchers*.

The authors thank our participants for their time and effort. This research was funded by the Israeli Science Foundation, Grant 1074/16

Aslib Journal of Information Management

2050-3806
DOI 10.1108/AJIM-08-2020-0250

The aims of this preliminary study are: (1) to determine whether there is a group of *hyper-searchers* who search significantly more than other people, (2) to study what drives *hyper-searchers* to search more, (3) to test which factors are associated with increased search behavior. Based on previous work that found linguistic processing is essential for search behavior, while visuospatial and place orientation brain regions are associated with navigation (Benn *et al.*, 2015a), we hypothesized that both high verbal memory and low visuospatial memory capabilities will have a positive effect on the tendency to search. In the current study, we asked 65 participants to estimate their percentage of search retrievals out of all file retrieval they conduct and then measured their verbal memory, visuospatial memory and several other factors. To determine whether there are *hyper-searchers*, we further re-analyzed the results from four previous studies, resulting in data from 1,317 participants (Bergman *et al.*, 2008, 2014, 2019a; Bergman and Yanai, 2018).

Theoretical background

Personal information management (PIM) is an activity in which individuals store and retrieve their personal information items. Such information items include files, emails, Web favorites, contacts and notes (Jones and Teevan, 2007; Bergman and Whittaker, 2016). Early PIM research tended to be qualitative and exploratory in nature, relying on interviews and observational studies (Kelly and Teevan, 2007). While such methods were important for identifying the basic PIM phenomena, as this research field matures, there has been a growing need for findings to be supported by more rigorous quantitative approaches that test statistical relations between variables (Bergman, 2013). The PIM subfield of file research includes studies on file *organization* (e.g. Dinneen *et al.*, 2019; Henderson and Srinivasan, 2009; Gonçalves and Jorge, 2003), file *sharing* (e.g. Rader, 2009; Berlin *et al.*, 1993; Voida *et al.*, 2006) and file *retrieval* (e.g. Bergman *et al.*, 2010, 2014, 2019a). This subfield is a rapidly growing area of research, as evidenced by a recently published review that identified over 230 publications on file research (Dinneen and Julien, 2020).

Search percentages

Retrieval is the main reason people manage their files. It is essential for retrievals to be both successful and efficient, because information cannot be used unless it can be re-accessed (Bergman and Whittaker, 2016). There are two main ways in which people retrieve their files: search and navigation. *Search* is a process in which users first generate a query specifying some attribute of the target item (e.g. its name or part of it, text it contains, its format, etc.). The search engine then returns a set of results from which the user selects the relevant item. By contrast, *navigation* is a folder-related process by which users manually traverse their folder hierarchy until they reach the folder in which the target item is stored, and then locate the target within that folder (Bergman *et al.*, 2008).

There seem to be clear advantages for search over navigation. Search is more flexible because unlike navigation, it does not require users to remember an exact folder location; instead, users can specify any attribute they happen to remember about the document in their query (Lansdale, 1988). Search can also be more efficient, as users can retrieve information by generating a single query, instead of using multiple operations to navigate through their folder hierarchy. Lastly, unlike navigation, search does not require users to engage in complex organizational strategies that exhaustively anticipate their future retrieval requirements, and therefore, it is claimed, to address the management problem (Cutrell *et al.*, 2006; Lansdale, 1988; Russell and Lawrence, 2007; Dourish *et al.*, 2000). These arguments against navigation are bolstered by historical developments in Web access. Here, the use of navigational systems such as Yahoo Web directories to retrieve Web data has largely disappeared, to be replaced by search (Kobayashi and Takeda, 2000; Obendorf *et al.*, 2007). These intuitive arguments have led many PIM researchers to propose that search

engines should replace folders (Cutrell *et al.*, 2006; Lansdale, 1988; Russell and Lawrence, 2007; Dourish *et al.*, 2000; Fertig *et al.*, 1996; Raskin, 2000).

However, in contrast to these claims, empirical research continues to show preference for navigation over search (Barreau and Nardi, 1995; Boardman and Sasse, 2004; Kirk *et al.*, 2006; Teevan *et al.*, 2004; Bergman *et al.*, 2014, 2019b; Bergman and Yanai, 2018). Moreover, since 2005, search engines began to use a constantly updated index system, which enabled newer search engines to run 1,000 times faster than old ones (Farina, 2005; Lowe, 2006) and consequently improve their interfaces. However, despite these improvements, a multi-method research, which combined a large-scale study of Mac search engines users and a longitudinal study of users of PC search engines, revealed no increases in search engine use among users of improved search engines compared to those who used old ones (Bergman *et al.*, 2008). Instead, these studies confirmed that search was utilized primarily as a *last resort* for a small minority of retrievals (an average search percentage of 4–15% depending on the study), when users could not recall the folder where they had stored their document. The same research also revealed no evidence that use of improved search engines led people to change their filing habits, such that they become less reliant on folders. A more recent logfile study testing a newer search engine validated these findings, confirming that participants used searches in just 4.5% of their document retrievals (Fitchett and Cockburn, 2015).

Why do people prefer navigation over search? A combination of cognitive (Bergman *et al.*, 2013) and neuroimaging (Benn *et al.*, 2015a) studies indicates that the folder navigation preference has deep neuro-cognitive roots: when navigating to their files, users employ the same brain structures that they use when navigating in the physical world. These same brain areas are also used for navigation by monkeys, rats and pigeons and require little verbal resources. By contrast, search recruits Broca's area, commonly observed in linguistic processing (e.g. Conner *et al.*, 2019). As such, given that most tasks we perform on computers require the use of language, searching may cause an interruption to the flow of work, while navigation would enable a person to continue to focus on their current document.

While navigation preference over search was repeatedly observed for the large majority of participants, there may be individual differences in the tendency to search. When reviewing data from previously published papers (Bergman *et al.*, 2008, 2014, 2019a; Bergman and Yanai, 2018), we noticed that search frequency has a long tail of people who search much more than others; we term these individuals *hyper-searchers*. The current study aims to explore whether there is indeed such a group, and if so, why do they search much more than other people?

Search tendency as a trait, and what influences it

Bergman and Yanai (2018) studied the retrieval of personal information across different formats (files and emails) and devices (personal computers and smartphones). Their results regarding the search percentage indicated both a between-subjects' variability and a within-subjects' consistency across different situations. In other words, while some participants tended to search more compared to others, the tendency was consistent and stable across different formats and devices. These results indicate that the tendency to search can be regarded as a personal trait.

If search tendency is a trait, then the next question would naturally be, what influences this trait? Bergman *et al.* (2019b) found that age is positively correlated with search percentage. A possible reason for this is that older people find it harder to remember where they located their files and are therefore forced to rely more heavily on retrieval by search. Similarly, Bergman *et al.* (2020) found that those who self-reported to have a better organized file system and to have a better memory for their file locations tended to search less. Interestingly, the search percentage did not correlate with the more general traits of being organized or having a good memory in general. A possible explanation for this is that there was a relatively low correlation between file organization as a specific behavior and organization as a general personality trait and between memory for files as a specific cognitive ability and memory as a general cognitive

trait. In other words, people can be organized in some situations (e.g. with their files) and unorganized in others (e.g. at home). Similarly, individuals can have a good memory for some things (e.g. files) but not for others (e.g. faces). This raises the question: what kind of more general memory traits affect search percentage? Two possible candidates for this are visuospatial memory, on which navigation in the real world depends, and verbal memory, as this may be needed for precise and accurate retrieval of search terms.

Research questions

(1) Is there a group of hyper-searchers who search significantly and substantially more than other people?

We defined *hyper-searchers* as participants that estimated their search percentage to be over one standard deviation above the average. We then compared their average search percentage with that of the other participants.

(2) What profile is associated with increased search behavior among *hyper-searchers?*

Following the neurocognitive results of Benn *et al.* (2015a) and Bergman *et al.* (2013), we have proposed two research hypotheses:

H1. *Hyper-searchers* have significantly better verbal memory than others, allowing them to recall search terms more accurately and easily than other participants. The accuracy and ease of search are important, as they compete with the navigation option, which was shown to be a highly effective and prevalent retrieval method (Bergman *et al.*, 2008, 2014, 2019a) that requires very little effort from the user (Bergman *et al.*, 2013). In addition, search requires the user to think of unique words within a file to avoid a long list of irrelevant files with similar names. Identifying such a unique term was previously shown to be difficult for users to identify in the context of online search (Benn *et al.*, 2015b).

H2. *Hyper-searchers* have significantly reduced visuospatial memory scores compared to other participants. Consequentially, they forget where they located their files and are therefore unable to navigate to them, resorting to search instead.

In addition, we compared *hyper-searchers* to other participants for a number of additional variables, including age, years of study, intelligence and computer literacy.

(3) Which cognitive traits are associated with increased search behavior?

Following the neurocognitive results of Benn *et al.* (2015a) and Bergman *et al.* (2013), we had two research hypotheses:

H3. Search percentage is positively correlated with verbal memory scores.

H4. Search percentage is negatively correlated with visuospatial memory scores (people with insufficient visual memory often fail to remember where they had located their files and turn to search instead).

We also tested the contribution of age, years of study, intelligence and computer literacy to search percentage.

Methodology

Research tools

Search percentages were evaluated in the questionnaire using the same question as in Bergman *et al.* (2008), as this method was found to be highly reliable when compared to

observation of search percentages during actual retrievals. Additional questions measured self-reported organizational skills and computer literacy (both on a 1–5 Likert scale), age, years of study and intelligence score if available (self-reported score in the psychometric test taken by university and college candidates, somewhat equivalent to the IQ test).

To examine the participants' verbal memory skills, we used the logical memory test from the Wechsler Memory Scale – Revised (WMS-R) (Wechsler, 1945). This is a test of verbal declarative memory, in which a story is read to the participant, and after the story, the participant recalls details from memory. Scores are given on the accuracy of the retelling of the story, with specific importance attributed to the use of exact and accurate use of words and terms. To reflect a memory task that is similar to file retrieval, we not only used the scores from logical memory II (LM-II), where recall was probed 30 min after the story was told, but also conducted LM-I, which involves immediate recall after the story is told, as this was part of the standard test procedure. We used the WMS-R because it tests the long-term memory recall of exact words. This is similar to search, which requires long-term memory recall of exact words from the file name or content. The use of "exact" terms and words is essential in this context, as using equivalent search terms does not yield the expected results. Two of the limitations of the WMS-R are that it is artificial and has little to do with the context of the participants' personal life, and that long-term memory is not tested beyond 30 min.

We used an online memory game to examine visuospatial memory [1]. It is an online version of the offline "Concentration" cards game [2]. The game presented face-down picture cards in random order so that participants could not see them. Using a computer mouse, participants had to click on a card to reveal its content. In each step, participants turned over two cards, which disappeared if their picture was identical (a pair), or turned back if their picture did not match. The goal of the game was to find all the pairs in as little steps as possible. To succeed, participants had to use their visual memory (to remember the pictures on the cards) and their spatial memory (to remember where the card was located). Therefore, the number of steps taken to complete the game measured their visuospatial memory on a reversed scale – the less steps it took them to finish the game, the better their visuospatial memory. In a pilot research, we found that 15 pairs were an optimal level of difficulty, which prevented both floor and ceiling effects. This task can be seen as equivalent to the task of navigating to a file that is hidden within the folder structure. As files are not visible to the users, the folders, which are located in different places in the hierarchy, must be opened to reveal their content. To do so, users recruit brain regions associated with visuospatial processing (Benn *et al.*, 2015a). Users choose the folder to be opened based on its location on the screen and the expected content it will reveal in the next step, which is equivalent to the process of successfully completing the online memory game in as few steps as possible (i.e. selecting a card based on its location on the screen and the expected content it will reveal once clicked). However, one limitation of this memory game is that unlike a file system, there are no "clues" as to the content of the card, in the same ways that there are clues about the content of a folder (e.g. folder name and/or its position in a hierarchical structure). Another limitation is that unlike a folder system, the user did not create the order of the cards, and hence may find it more difficult to remember the locations of items (as the learning is done through exploration).

Participants

We recruited 65 Israeli participants using convenience sampling based on personal acquaintance. Of the participants, 46 (71%) were women. Their ages ranged from 27 to 67 years old ($M = 42.02$, SD = 11.95), level of education ranged from 12 to 25 years of education ($M = 16.5$, SD = 2.74) and computer literacy was high ($M = 4.10$ out of 5, SD = 0.76).

Procedure

The tester explained the study to the potential participants and asked them to sign an informed consent form approved by our university's ethics committee. The experiment began with the Wechsler memory test. The tester read the story slowly and clearly. To complete LM-I, participants were then asked to retell the story with as many details as possible. As instructed by the test manual, the tester kept a blank expression on her face while participants retold the story and did not answer any questions. Next, participants were asked to play the online memory game, Screen-recording software (Icecream Screen Recorder) was used to record the game, and the tester noted the number of steps it took each participant to find all the pairs. Following the visuospatial and the verbal memory tests, the tester and participant filled out the questionnaire together. Once 30 min had elapsed since the LM-I story was retold, participants were asked to retell the story again with as many details as they could recall for the LM-II test. Again, the tester marked all the details that were related. Responses were recorded for later reliability testing by the research team. The tester then thanked the participants and paid them for their time.

Post-study

Following the analysis of results, and the identification of a group of *hyper-searchers,* the tester returned to each of these participants for further testing. Specifically, the tester measured the level of file organization using a "guided tour" technique in which the participant showed the tester how their files are organized, and by the participant using a 1–5 Likert scale. Next, a semi-structured interview was conducted, with the participants asked to elaborate on aspects such as how they decide what retrieval method to use to retrieve a file; when do they tend to use search and whether they use search only as a last resort.

Results

The average search percentage by the participants was 15% of all file retrievals, with a standard deviation of 16%.

There is a small group of people with a significantly and substantially higher search percentage

RQ1. Is there a group of hyper-searchers that search significantly and substantially more than other people?

We defined *hyper-searchers* as participants whose estimated search percentage was over one standard deviation above the sample mean (i.e. above 31%). Of the 65 participants, seven (11%) met this criterion. The average search percentage for the seven *hyper-searchers* was 51% (SD = 14%), which was more than five times that of the other 58 participants ($M = 10\%$, SD = 9%). An independent *t*-test indicated that the difference between the groups was significant $t(63) = 10.72$, $p < 0.001$.

To validate this finding, we conducted new analyses of results using data from previous published work (Table 1). These include (Bergman *et al.*, 2008), in which participants self-reported their search percentage, as in the current study, and (Bergman *et al.*, 2014, 2019a; Bergman and Yanai, 2018), in which we observed the participants' retrieval behavior.

The results presented in Table 1 validate the existence and stable prevalence of *hyper-searchers*. In particular, note the consistency between results in: (1) the cutting point between the *hyper-searcher* group and the other participants (30–31%), (2) the percentage of *hyper-searchers* among all participants (8–13%) and (3) the search percentage among

Method of data collection		Current paper Self-report	Bergman *et al.* (2008) Self-report	Bergman *et al.* (2014) Observed behavior	Bergman and Yanai (2018) Observed behavior	Bergman *et al.* (2019) Observed behavior
All participants	*N*	65	628	275	60	289
	Mean search	15%	13%	8%	6%	12%
	SD	16%	17%	23%	15%	21%
Cutting point between groups		31%	30%	31%	31%	31%
Hyper-searchers	*N* (%)	7 (11%)	80 (13%)	34 (12%)	5 (8%)	37 (13%)
	Mean search	51%	51%	64%	43%	59%
	SD	14%	19%	24%	19%	5%
Other participants	Mean search	10%	7%	1%	2%	5%
	SD	9%	7%	4%	8%	10%
p of *t*-test		<0.001	<0.001	<0.001	<0.001	<0.001

Table 1.
Prevalence of *hyper-searchers* in the current study and re-analyzing older data (N = 1,317)

hyper-searchers (43–64%). Note also that in all studies, the *hyper-searchers* conducted searches several times more than the other participants. The accumulation of these results indicates that there is a group of *hyper-searchers* who search for their files significantly and substantially more than the rest of the population.

Hyper-searchers *have better verbal memory than the rest of the participants*

RQ2. What profile is associated with the increased search behavior among *hyper-searchers?*

To test our first hypothesis that *hyper-searchers* have better verbal memory, we conducted an independent *t*-test on the WMS-R LM-II test (memory recall after a 30 min delay). The results indicated that the verbal memory score of the *hyper-searchers* ($M = 14.86$, SD = 2.87) was significantly higher than that of the rest of the participants ($M = 12.02$, SD = 3.59), $t(63) = 2.01$, $p < 0.01$. This suggests that *hyper-searchers* have better verbal memory than the rest of the population, which is likely to enable them to remember exact words contained in files, and then search for them more easily and successfully than the rest of the population.

To test our second hypothesis that *hyper-searchers* have reduced visual memory performance, we compared the number of steps needed to complete a visuospatial memory-game task using an independent *t*-test. No significant result was found between *hyper-searchers* ($M = 36.43$, SD = 8.50) and the rest of the participants ($M = 36.21$, SD = 7.67).

We also compared *hyper-searchers* to our other participants on a number of other variables, including age, years of study, intelligence, self-reported organizational skills and computer literacy by using independent *t*-tests. No significant results were found. In addition, we used the results of Paper 3 to test whether *hyper-searchers* were different than the other participants in the big five personality traits test. Again, no significant results were found.

Lastly, we conducted a post-test with the *hyper-searchers* to test their file organization. When asked "How well are your files organized?," none of the participants reported a disorganized file hierarchy (two participants answered "to a medium degree," four "to a high degree" and one "to a very high degree"). In addition, according to the interviewer, when conducting a guided tour of their file system, none of the participants displayed an unorganized file hierarchy. Importantly, six out of seven *hyper-searchers* insisted that they search for their files only as a last resort, typically after they failed to navigate to the file.

Verbal memory and visuospatial memory predict search percentage conversely

RQ3. Which cognitive traits are associated with increased search behavior?

We conducted a simultaneous multiple linear regression analysis to test the effect of verbal memory, visuospatial memory, age, years of study, intelligence, order (tendency to organize) and computer literacy on search percentage. The regression model is presented in Table 2.

Table 2 indicates that both verbal and visuospatial memory significantly predict search percentage of files.

Verbal memory effect. The regression model indicates that verbal memory positively predicts search percentage. Note that this result is not independent of the one presented in response to RQ2, but rather a different analysis of the same phenomena. Figure 1 is a scatter plot representing verbal memory scores and search percentage.

Visuospatial memory effect. The explanation of this effect is a bit more complex. Visuospatial memory is presented in a reverse scale (the more steps it takes to complete the memory game task, the worse the participants' visuospatial memory). Our results indicate

that the worse our participants' visuospatial memory is, the more they resort to search as a retrieval method. This is likely to be due to them forgetting where their files were, hence being unable to retrieve the file using navigation. Figure 2 is a scatter plot of number of steps in card game and search percentage.

Discussion

Research consistently indicates navigation preference over search (Barreau and Nardi, 1995; Boardman and Sasse, 2004; Kirk *et al.*, 2006; Teevan *et al.*, 2004; Bergman *et al.*, 2014, 2019b; Bergman and Yanai, 2018). The results from this study reaffirm this assertion, by showing that the average search percentage remains around 15%. However, this study furthers our understanding of PIM behavior, by revealing that there is a small but consistent group of people (11% in the current study), whom we have termed *hyper-searchers*. These individuals search significantly more than others. The seven *hyper-searchers* in the current study searched in 51% of their retrievals, which is about five times more than the other participants (10% on average). We have been able to validate these results by re-analyzing data from previous papers that measured search percentage using both participants' estimations (Bergman *et al.*, 2008) and observed participants' actual retrieval behavior (Bergman *et al.*, 2014, 2019a; Bergman and Yanai, 2018). The weighted average of these four studies with a

Parameter	Coefficient	Standardized coefficient	*SE*	*t*	*p*
Constant	−15.20		40.11	−0.38	0.71
Verbal memory	2.33	0.44	0.96	2.42	0.02*
Visuospatial memory (reverse scale)	1.60	0.57	0.50	3.18	0.004**
Age	−0.48	−0.20	0.46	−1.07	0.30
Order	−7.56	−0.34	3.89	−1.94	0.06
Intelligence	−0.002	−0.01	0.04	−0.06	0.95
Years of education	−1.53	−0.17	1.62	−0.94	0.35
Computer literacy	5.56	0.27	3.54	1.57	0.13

Note(s): *$p < 0.05$; **$p < 0.01$

Table 2.
Regression model predicting search percentage ($R^2 = 0.45$, $p < 0.05$)

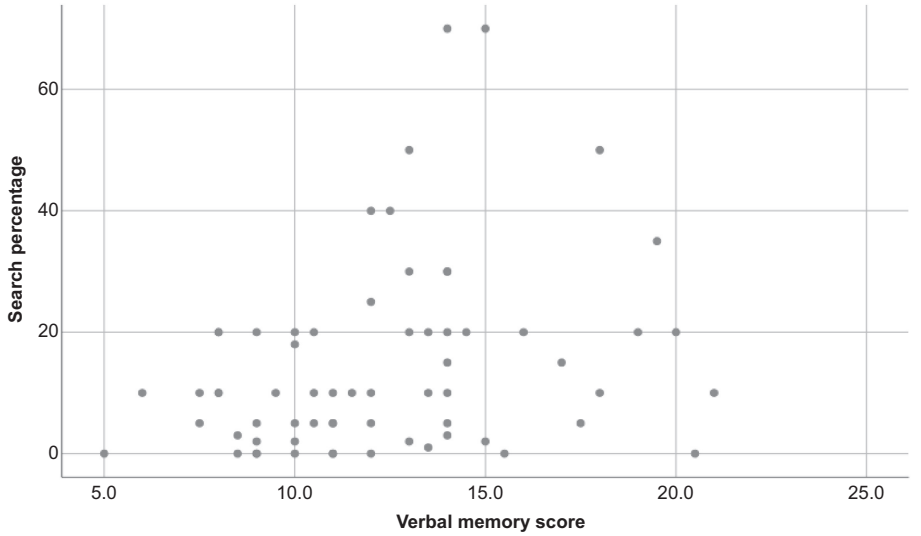

Figure 1.
Scatter plot of verbal memory and search percentage

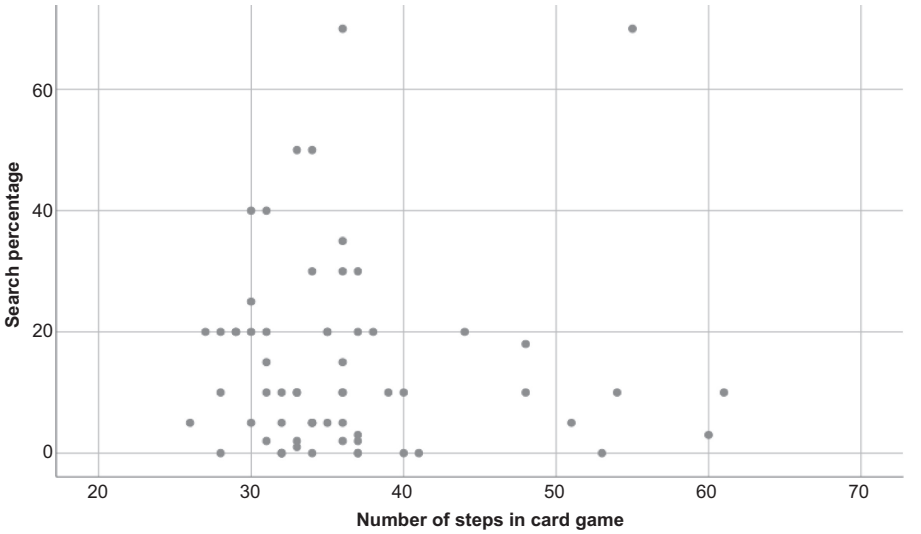

Figure 2.
Scatter plot of number of steps in card game and search percentage

total of 1,252 participants revealed very similar results: The *hyper-searchers* amount to 11% of the participants, and their average search percentage was 55% – 11 times more than the average of other participants (5%). Combined, these results provide strong evidence for constructing a group of *hyper-searchers*.

The current study also found that *hyper-searchers* had significantly better verbal memory compared to the other participants, and that the search percentage in general was positively predicted by verbal memory scores and negatively predicted by visuospatial memory scores. These will be discussed in the next two sections.

Verbal memory

Despite some limitations, the logical memory subtest of the WMS is commonly employed to assess immediate and delayed (30 min) verbal retention (Gass *et al.*, 2020). The particular focus of the test on recall of exact words, rather than general context, was of particular importance to us in this study, as desktop search requires the exact term to be generated, often after a delay period in which the user had not engaged with the file.

Previous work revealed that engaging the language faculty for a shadowing task disrupts searching but not navigation (Bergman *et al.*, 2013). Furthermore, search (but not navigation) recruits Broca's area (Benn *et al.*, 2015a), which has recently been shown to be particularly involved in word retrieval and selection (Conner *et al.*, 2019). The current findings, therefore, add to this previous body of work by demonstrating that those who score higher on the verbal logical memory delayed recall test are significantly more likely to use search. While most of the *hyper-searchers* in the current study claimed to use search only when navigation failed, we propose that the ease by which they are able to generate an exact word enables them to resort to search faster (rather than try another navigation attempt) than the rest of the participants, and in some cases (as reported by one of the *hyper-searchers*) even to use search as first retrieval choice, rather than as a "last resort."

Visuospatial memory

Poor visuospatial memory was a significant predictor of search percentage in the regression model. Previous work suggested that navigation relies on visuospatial processing in the brain

(Benn *et al.*, 2015a), and it is, therefore, likely that those with poor visuospatial processing find it more difficult to remember where their files are located forcing them to resort to search. Further support for the idea that poorer visuospatial memory may lead to increase in the "fall back" method of search was found by Bergman *et al.* (2019b) who showed that search percentage increases with age, possibly because of a deterioration in visuospatial memory (Flicker *et al.*, 1984).

One practical implication of this study is the potential to improve the interface of file retrievals, to provide specific support for verbal and visuospatial memory for users who present with difficulties in one or both areas.

Limitations and future research

This study suffers from the following limitations that can be addressed in future research:

Search percentage measure limitations: The results regarding the existence of *hyper-searchers* group seems strong as they are confirmed by re-analyses of our previous studies (Bergman *et al.*, 2008, 2014, 2019a; Bergman and Yanai, 2018). However, all of these studies used *either* estimations or observations of a small number of retrievals per participant, and both methods have their limitations. Although estimates were shown to be highly reliable in general (Bergman *et al.*, 2008), a small number of participants may have grossly over-estimated their search percentage. With regard to observations, the high search percentage found among a small number of participants could alternatively be explained by target files that are particularly difficult to find using navigation (e.g. because these files have not been retrieved for a long time), forcing participants to resort to search. In a future study, it would be good to use a double screening procedure in which participants assigned to the *hyper-searchers* group will have particularly high scores in *both* estimated and observed search percentages, reducing the risk of incidental occurrence of high-search scores or self-reporting errors.

Small number of hyper-searchers: Only seven *hyper-searchers* were examined in this preliminary study; therefore, the results regarding the reasons for their extensive search behavior should be treated with caution. In addition, we did not test for other possible reasons, such as their past experience with file search engines. Future research should address these shortcomings.

Future studies should also test whether *hyper-searchers* search more successfully and efficiently better than other people, and if so, explore what we can learn from them to improve file searching by other users.

Notes

1. https://www.helpfulgames.com/subjects/brain-training/memory.html
2. https://en.wikipedia.org/wiki/Concentration_(card_game)

References

Barreau, D.K. and Nardi, B.A. (1995), "Finding and reminding: file organization from the desktop", *SIGCHI Bulletin*, Vol. 27 No. 3, pp. 39-43.

Benn, Y., Bergman, O., Glazer, L., Arent, P., Wilkinson, I.D., Varley, R. and Whittaker, S. (2015a), "Navigating through digital folders uses the same brain structures as real world navigation", *Scientific Reports*, Vol. 5, pp. 1-8.

Benn, Y., Webb, T.L., Chang, B.P. and Reidy, J. (2015b), "What information do consumers consider, and how do they look for it, when shopping for groceries online?", *Appetite*, Vol. 89, pp. 265-273.

Bergman, O. (2013), "Variables for personal information management research", *ASLIB Proceedings*, Vol. 65 No. 5, pp. 464-483.

Bergman, O. and Whittaker, S. (2016), *The Science of Managing Our Digital Stuff*, Mit Press, Cambridge, MA.

Bergman, O. and Yanai, N. (2018), "Personal information retrieval: smartphones vs. computers, emails vs. files", *Personal and Ubiquitous Computing*, Vol. 22 No. 4, pp. 621-632.

Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N. and Whittaker, S. (2008), "Improved search engines and navigation preference in personal information management", *ACM Transactions on Information Systems*, Vol. 26 No. 4, pp. 1-24.

Bergman, O., Whittaker, S., Sanderson, M., Nachmias, R. and Ramamoorthy, A. (2010), "The effect of folder structure on personal file navigation", *Journal of the American Society for Information Science and Technology*, Vol. 61 No. 12, pp. 2426-2441.

Bergman, O., Tene-Rubinstein, M. and Shalom, J. (2013), "The use of attention resources in navigation vs. search", *Personal and Ubiquitous Computing*, Vol. 17 No. 3, pp. 583-590.

Bergman, O., Whittaker, S. and Falk, N. (2014), "Shared files – the retrieval perspective", *Journal of the American Society for Information Science and Technology*, Vol. 65 No. 10, pp. 1949-1963.

Bergman, O., Israeli, T. and Whittaker, S. (2019a), "Factors hindering shared files retrieval", *Aslib Journal of Information Management*, Vol. 72 No. 1, pp. 130-147.

Bergman, O., Israeli, T. and Whittaker, S. (2019b), "Search is the future? The young search less for files", *Proceedings of the Association for Information Science and Technology*, Vol. 56 No. 1, pp. 360-363.

Bergman, O., Israeli, T., Whittaker, S., Yanai, N. and Amichai-Hamburger, Y. (2020), "The effect of personality traits on file retrieval", *iConference 2020*, pp. 1-8.

Berlin, L.M., Jeffries, R., O'Day, V.L., Paepcke, A. and Wharton, C. (1993), "Where did you put it? Issues in the design and use of a group memory", *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems*, New York, pp. 23-30.

Boardman, R. and Sasse, M.A. (2004), "'Stuff goes into the computer and doesn't come out': a cross-tool study of personal information management", in *SIGCHI Conference on Human Factors in Computing Systems*, Vienna, pp. 583-590.

Conner, C.R., Kadipasaoglu, C.M., Shouval, H.Z., Hickok, G. and Tandon, N. (2019), "Network dynamics of Broca's area during word selection", *PloS One*, Vol. 14 No. 12, p. e0225756.

Cutrell, E., Dumais, S.T. and Teevan, J. (2006), "Searching to eliminate personal information management", *Communications of the ACM*, Vol. 49 No. 1, pp. 58-64.

Dinneen, J.D. and Julien, C.A. (2020), "The ubiquitous digital file: a review of file management research", *Journal of the Association for Information Science and Technology*, Vol. 71 No. 1, pp. E1-E32.

Dinneen, J.D., Julien, C.-A. and Frissen, I. (2019), "The scale and structure of personal file collections", *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 327.

Dourish, P., Edwards, W.K., LaMarca, A., Lamping, J., Petersen, K., Salisbury, M., Terry, D.B. and Thornton, J. (2000), "Extending document management systems with user-specific active properties", *ACM Transactions on Information Systems*, Vol. 18 No. 2, pp. 140-170.

Farina, P.A. (2005), *A Comparison of Two Desktop Search Engines: Google Desktop Search (Beta) vs. Windows XP Search Companion*, 21st Computer Science Seminar, Hartford.

Fertig, S., Freeman, E. and Gelernter, D. (1996), "'Finding and reminding' reconsidered", *SIGCHI Bulletin*, Vol. 28 No. 1, pp. 66-69.

Fitchett, S. and Cockburn, A. (2015), "An empirical characterisation of file retrieval", *International Journal of Human-Computer Studies*, Vol. 74, pp. 1-13.

Flicker, C., Bartus, R.T., Crook, T.H. and Ferris, S.H. (1984), "Effects of aging and dementia upon recent visuospatial memory", *Neurobiology of Aging*, Vol. 5 No. 4, pp. 275-283.

Gass, C.S., Patten, B., Penate, A. and Rhodes, A. (2020), "An enhanced delayed recognition measure for the Logical Memory subtest of the Wechsler Memory Scale–IV", *Applied Neuropsychology: Adult*, pp. 1-5.

Gonçalves, D. and Jorge, J.A. (2003), "An empirical study of personal document spaces", *Proceedings of DSV-IS'03*, Springer, Berlin and Heidelberg, pp. 46-60.

Henderson, S. and Srinivasan, A. (2009), "An empirical analysis of personal digital document structures", *HCI International 2009*, Springer, Berlin Heidelberg, pp. 394-403.

Jones, W. and Teevan, J. (2007), *Personal Information Management*, University of Washington Press, Seattle, WA.

Kelly, D. and Teevan, J. (2007), "Understanding what works: evaluating PIM tools", in William, P.J. and Teevan, J. (Eds), *Personal Information Management*, University of Washington Press, Seattle, pp. 190-204.

Kirk, D., Sellen, A., Rother, C. and Wood, K. (2006), "Understanding photowork", *SIGCHI Conference on Human Factors in Computing Systems*, ACM, New York, pp. 761-770.

Kobayashi, M. and Takeda, K. (2000), "Information retrieval on the web", *ACM Computing Surveys*, Vol. 32 No. 2, pp. 144-173.

Lansdale, M.W. (1988), "The psychology of personal information management", *Applied Ergonomics*, Vol. 19 No. 1, pp. 55-66.

Lowe, M. (2006), *Evaluation of Desktop Search Applications*, Kalio, Sydney.

Obendorf, H., Weinreich, H., Herder, E. and Mayer, M. (2007), "Web page revisitation revisited: implications of a long-term click-stream study of browser usage", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, New Yorks, pp. 597-606.

Rader, E. (2009), "Yours, mine and (not) ours: social influences on group information repositories", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, New Yorks, pp. 2095-2098.

Raskin, J. (2000), *The Humane Interface: New Directions for Designing Interactive Systems*, ACM Press/ Addison-Wesley Publishing Co., Boston.

Russell, D. and Lawrence, S. (2007), "Search everything", in Jones, W. and Teevan, J. (Eds), *Personal Information Management*, University of Washington Press, Seattle and London, pp. 153-166.

Teevan, J., Alvarado, C., Ackerman, M.S. and Karger, D.R. (2004), "The perfect search engine is not enough: a study of orienteering behavior in directed search", in Dykstra-Erickson, E. and Tscheligi, M. (Eds), *SIGCHI Conference on Human Factors in Computing Systems*, ACM Press, New York, pp. 415-422.

Voida, S., Edwards, W.K., Newman, M.W., Grinter, R.E. and Ducheneaut, N. (2006), "Share and share alike: exploring the user interface affordances of file sharing", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, New York, pp. 221-230.

Wechsler, D. (1945), *The Wechsler Memory Scale*, The Psychological Corporation, San Antonio, TX.

Corresponding author
Ofer Bergman can be contacted at: oferbergman@gmail.com